Modeling PRRSV-2 prevalence within a farrowing room

Onyekachukwu H. Osemeke[1], Eduardo de Freitas Costa[2], Vinicius Weide[3], Swaminathan Jayaraman[1], Gustavo S. Silva[1], Daniel C. L. Linhares[1]

[1]Veterinary Diagnostic and Production Animal Medicine, College of Veterinary Medicine, Iowa State University, Ames, Iowa, USA.

[2]Department of Epidemiology, Bioinformatics and Animal Models, Wageningen Bio-veterinary Research, Lelystad, The Netherlands.

[3]Instituto Federal Rio Grande do Sul

Instituto Federal de Educação, Ciência e Tecnologia do Rio Grande do Sul, Farroupilha, RS, Brazil

e-mail: vinicius.rodrigues@farroupilha.ifrs.edu.br

ABSTRACT

Commented [dFCE1]: Henry, please, do not use reference in the abstract

Estimated disease prevalence is an important assumption in calculating the minimum number of animals/units to be sampled for PRRSV surveillance/monitoring programs (Christopher-Hennings et al., 2001; Holtkamp et al., 2011, 2021; Rotolo et al., 2017; Rovira et al., 2007); these programs consider the minimum proportion of animals within a herd that is likely positive (viremic) for PRRSV to estimate the number of animals to be sampled. Family oral fluids (FOFs) have been demonstrated to be a more practical sample type for low-prevalence PRRSV surveillance in weaning-age pigs (Almeida, Zhang, Zimmerman, et al., 2021; Osemeke et al., 2022), since this sample is an aggregate sample type collected at the level of the crate/litter, the minimum number of litters/units to be sampled (sample size) will depend on the estimated number of positive litters/crates in a farrowing room, the diagnostic performance of the test, and a chosen confidence level (Osemeke et al., 2022; Rotolo et al., 2017).

This study aimed to characterize a relationship between the proportion of PRRSV-positive (viremic) pigs, the proportion of PRRSV-positive litters, and the likely proportion of litters to be positive by a FOF RT-rtPCR test in a farrowing room, using Monte Carlo simulations and data from an earlier study (Almeida, Zhang, Lopez, et al., 2021). Expectedly, the spatial distribution (clustering parameter) of viremic pigs within a farrowing room influenced the proportion of litters having at least one viremic piglet, and the number of positive FOFs expected from the room. This study assessed and measured this clustering parameter from the reference study and applied

I think you should summarize this paragraph in 3-4 lines, then go straight to "this study (...)" line 23.

Reason: I believe these informations should be placed in "Introduction" rather than in the Abstract.

different values of the clustering parameter to the stochastic models to estimate a relationship between all three proportions.

The results of this study provide guidelines on how practitioners can ascertain the number of FOF samples needed to match serum samples at an assumed piglet-level prevalence and clustering level. The results from this study also provide a platform for swine practitioners to estimate the likely proportion of viremic pigs in their barns, given the PRRSV-RT-rtPCR positivity rate of FOF samples submitted from a farrowing room.

From our findings, an estimated level of clustering of viremic pigs is worth considering when estimating the litter-level prevalence of PRRSV for onward FOF sample size calculations.

Introduction

Porcine respiratory and reproductive syndrome virus (PRRSV) poses a significant challenge to the global swine industry (Calderón Díaz et al., 2020; Holtkamp et al., 2013). Monitoring/surveillance remains an integral component of PRRSV control and elimination programs, and ascertaining the PRRSV status of pig populations around the time of weaning is crucial to guide decisions on health interventions and pig flow (Holtkamp et al., 2011).

Efficient PRRSV surveillance/monitoring programs allow for the early detection of infection and helps evaluate changes in PRRSV prevalence over time; aiding swine producers and veterinarians alike to forestall the spread of PRRSV (Mccaw, 2000; Silva et al., 2017), and evaluate progress made with instituted PRRSV management programs (Holtkamp et al., 2021; Linhares et al., 2014).

Different sample types are routinely submitted to Veterinary Diagnostic Laboratories in the US for PRRSV investigation by RT-rtPCR; these would include samples taken from individual pigs such as serum, swabs, semen, and post-mortem tissues; or aggregate samples taken from multiple pigs such as processing fluids and oral fluids (Trevisan et al., 2019). These samples are either submitted and tested individually or in pools.

The number of samples submitted for disease pathogen investigation is crucial to the success of a surveillance/monitoring exercise. Guided by epidemiological/statistical assumptions, the sample size should have enough power to detect at least one positive unit if the herd is truly positive for the pathogen of interest (Cameron et al., 2020; Stevenson, 2021).

Estimated prevalence at the individual pig level is one of the key variables used in calculating sample size to demonstrate disease freedom (Fosgate, 2009; Stevenson, 2021).

The preferred sample for PRRSV surveillance in sow herds is serum from weaning-age pigs (Holtkamp et al., 2011). Although the serum sample is the sample of choice for PRRSV surveillance, it requires more skill, more manpower, is less animal welfare friendly, and is often impractical for frequent PRRSV surveillance in large herds (Turlewicz-Podbielska et al., 2020) compared to other (alternative) sampling options. For these reasons, since 2018, validated aggregate alternative samples have been the most frequently submitted samples for PRRSV surveillance in the US (Trevisan et al., 2019).

~~Almeida (~~Nunes de Almeida~~,~~ (2020) demonstrated that, especially in low prevalence, ~~Family~~ family oral fluids (FOFs) are a more convenient and cost-efficient alternative to serum sampling for PRRSV surveillance in weaning-age pigs. A FOF sample is an aggregate sample obtained when oral fluids are wrung off a rope chewed by a sow and her piglets (Almeida et al., 2020). A challenge with interpreting a positive result from testing FOF and other aggregate sample types is that one only knows that at least one animal that contributed to the sample matrix is pathogen-positive (shedding) but cannot ascertain the exact number of animals shedding.

Consequently, little to nothing is known about the number of positive pigs in a sampled room, given the proportion of PRRSV-positive aggregate samples, such as FOF, obtained from that room.

The individual pig is the unit for which sample size is calculated when non-aggregate samples are collected, while the litter is the unit for which sample size is calculated when an aggregate sample such as FOF is to be collected (Osemeke et al., 2022; Rotolo et al., 2017) it will be helpful to swine practitioners in making sampling decisions if they understood how the proportion of PRRSV-positive (viremic) piglets related with the proportion of PRRSV-positive litters, as both parameters are needed assumptions in estimating sample sizes.

To the best of our knowledge, the relationship between the mentioned proportions in a swine farrowing room has not been previously characterized. This study builds upon a previous study that assessed the observed probability of a positive FOF sample given the number of viremic PRRSV RT-rtPCR positive piglets within a litter (Almeida, Zhang, Lopez, et al., 2021)~~,~~. We further developed an *in-silico* study to assess and model a relationship between the ~~Piglet~~ piglet level prevalence (PP), ~~True~~ true litter-level PRRSV prevalence (TLP) and ~~Apparent~~ apparent litter-level PRRSV prevalence (ALP) in a farrowing room. It is expected that the results from this study will not only provide insights to swine practitioners as to the likely relationship between PP, TLP, and ALP, but also provide a template/framework for determining the piglet-level prevalence of PRRSV (and potentially other swine pathogens) that are being monitored using aggregate samples.

2. Methodology

2.1 PRRSV detection in pig litters using FOF

Based on a dataset from ~~Almeida et al. (~~Almeida, Zhang, Zimmerman, et al.~~,~~ (2021) 199 litters had all piglets sampled for PRRSV RNA detection by RT-rtPCR (reverse transcription real-time polymerase chain reaction); also, each litter (i=1,...,199) was sampled using FOF. The litters were sampled from 11 farrowing rooms across six different swine breeding farms (j=1,...,6).

The effect of the proportion of positive piglets (x) on the detection of a positive litter using FOF (P^{FOF}) was assessed with a generalized linear mixed model employing a logit link function and a 'residual' Bernoulli distribution. In addition, the linear predictor comprised random effects for farms according to:

$$logit(P_{ij}^{FOF}) = \alpha + \beta x_{ij} + \varepsilon_{ij} + \gamma_j, \quad (1)$$

where α is the intercept of the model. ε_i is the random error assumed $\varepsilon_{ij} \sim N(0, \sigma_\varepsilon)$, and γ_j is the random effect accounting for the farm-effect in the model, assumed $\gamma_j \sim N(0, \sigma_\gamma)$, where ε_{ij} *and* γ_j are all independent. Approximate maximum likelihood inference was based upon Laplacian integration, as implemented in R (R Core Team, 2018) in routine *glmer* from library *lme4* (Bates et al., 2015).

Replace "." with ","

2.2 Stochastic model

I'm not sure, but I believe the expression " let's" doesn't match the rest of the text, where the third person and passive voice are used.

Suggestion: A random variable is considered (...)

Let's consider a random variable: number of positive piglets in the *i*-th litter (N_i), assuming that each piglet's status (positive/negative) is a Bernoulli trial, with a fixed *p* probability, thus N_i arises from a binomial process. Consider a room with *n* litters with different sizes (T_i) drawn from a discrete empirical distribution, and total number of piglets in the room $T = \sum_{i=1}^{n} T_i$. In a simplistic scenario, the allocation of positive piglets in each litter (N_i) would follow the relative size of the litter in the room. However, given that we are modeling an infectious disease, there might be situations where the total number of positive animals (*N*) may be "clustered" in a few litters (Carpenter, 2001; Kostoulas et al., 2013).

Accounting for this, the number of positive animals in each *i* litter (N_i) is calculated as a special case of the multinomial distribution, sampling recursively from binomial distributions using a clustering factor:

$$N_i|N_j(j = 1, \dots, i-1) \sim min\left\{Bin\left[\left(N - \sum_{j=1}^{i-1} N_j, pl_i\right)\right], \mathrm{T_i}\right\}, \quad (2)$$

where *j* stands for the successive allocation of positive animals within each litter. pl is the probability of success in this binomial process. Finally, *pl* is defined as:

and pl is (...)

binomial process, defined as:

$$pl_i = \frac{\mathrm{T_i}}{T - \sum_{j=1}^{i-1} T_j} + \left(1 - \frac{\mathrm{T_i}}{T - \sum_{j=1}^{i-1} T_j}\right) \cdot c, \; c \in [0,1]. \quad (3)$$

The reader should be aware that *c* is a clustering factor. Thus when $\lim_{c \to 1} pl = 1$, the positive piglets will be totally clustered in the smallest number of litters as possible. On the other hand, when $\lim_{c \to 0} pl = \frac{\mathrm{T_i}}{T - \sum_{j=1}^{i-1} T_j}$, piglets will be spread according to the relative size (number of piglets) of each litter regarding the room size.

To obtain the baseline clustering factor *c* we used the observed distribution of the within litter prevalence (θ_i) reported in ~~Almeida et al. (~~Almeida, Zhang, Zimmerman, et al.~~,~~ (2021) across seven rooms, each room with *n* litters. The lost function was the minimization of the mean squared errors of the predicted (eq 2) *vs* observed distribution of the within litter prevalence $f(c, \theta) = \frac{1}{n}\sum_{i=1}^{n}(\theta_i - \frac{N_i}{T_i})^2$. The objective function $f(c, \theta)$ can be used to calculate a parameter estimate $\hat{c} = argmin\left(f(c, \theta)\right)$. Each room was randomly chosen, 10000 times obtaining the parameters θ_i, N_i, and T_i. For each sampled room, 1000 acquisition points in the parameter space of *c* were sampled from a uniform distribution $c \sim uniform(0, 1)$, obtaining a distribution to optimize $\hat{c}$.

observed distribution was used...

2.2.2 *Apparent prevalence at the litter level*

The simulated proportions of positive piglets per litter obtained from 2.2 were used as input for the logistic model fit in 2.1, calculating the probability of detection of each simulated litter using FOF sampling. Now we are interested in modeling a random variable (*S*) describing the most probable number of positive litters detected in a routine FOF sampling in a farrowing room. Assuming the probability of each litter being detected by FOF (P_i^{FOF} see eq. 1) are independent of each other, and the positive/negative status of a litter $y_i \sim Bernoulli(P_i^{FOF})$, S equals $\sum_{i=1}^{k} y_i$. The expected apparent litter prevalence (ALP) was obtained as S/n. S was generated a total of 2,000 times to improve the accuracy of the Monte Carlo estimation, and the mean ALP was obtained and stored for that iterated room.

Commented [dFCE2]: You need to explain why to run it for 2000, or wherever the number of times.

Your readers need to understand, why to do this

The parameters and distributions used in the simulations are described in table 1. In this simulation, 5000 stochastic iterations were performed, each one representing a different room, propagating the between litter variability observable in different farrowing rooms.

Table 1 – Descriptions of baseline model parameters used to compare the true and apparent liter prevalence of PRRSV.

Parameter/variable	Distribution/function	Description	Source
p	Fixed= (range of values from 1% to 100%)	Probability of a piglet being positive in a room (prevalence)	Almeida
N	$p \cdot T$	Total number of positive animals in the room	Calculation
T	$\sum_{i=1}^{n} T_i$	Total number of piglets in the room	Calculation
T_i	empirical {(), ()} *	Number of piglets in the *i*-th litter	Almeida
n	Fixed=56	Number of crates in a room	Authors' opinion
N_i	$min\left\{Bin\left[\left(N - \sum_{j=1}^{i-1} N_j, pl\right)\right], T_i\right\}$	Number of positive piglets in *i*-th litter	Calculation

pl_i	$\frac{T_i}{T-\sum_{j=1}^{i-1} T_j} + \left(1 - \frac{T_i}{T-\sum_{j=1}^{i-1} T_j}\right) \cdot c$	Probability of success in this binomial process (i.e., allocation of positive piglets in a litter) for the in *i*-th litter	Calculation
c	Fixed=0.61	Clustering factor	Optimized based on Almeida

* empirical {(3, 4, 5, 6, 7, 8, 9, 10 ,11, 12, 13, 14, 15, 25), (0.0092, 0.0092, 0.0046, 0.0046, 0.0553, 0.0691, 0.0922, 0.1014 0.1982, 0.2074, 0.1244, 0.0783, 0.0415, 0.0046)}

2.3 Sensitivity analysis

To assess the effect of the clustering factor (*c*) and the room size (*n*) on the estimated relationship between pig-level-prevalence and litter-level prevalence we selected five values for *c* (0.05, 0.34, 0.63, 0.83, 1) and five values for *n* (10, 33, 56, 79, 102) combining them as a factorial design for the sensitivity analysis, totaling 25 different scenarios.

3. Results

3.1 PRRSV detection in pig litters using FOF

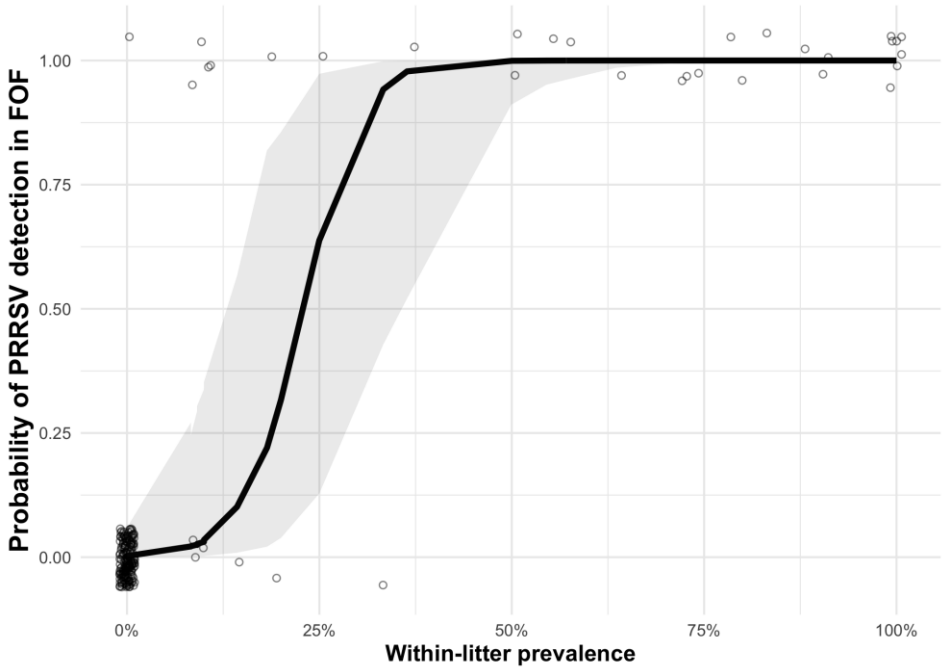


Figure 1: A jitter plot of the Probability of PRRSV detection in FOF by the proportion of positive pigs within litters (within litter prevalence). 95% prediction intervals are represented by the grey region around the regression line.

3.2 Stochastic model

3.2.1 Observed distribution of clustering in sampled farms

The clustering distribution across all sampled rooms had a minimum value of 0.00136, a median of 0.61, a mean of 0.57, and a maximum value of 1. The distributions of the clustering parameter across all sampled rooms are represented in figure 2.

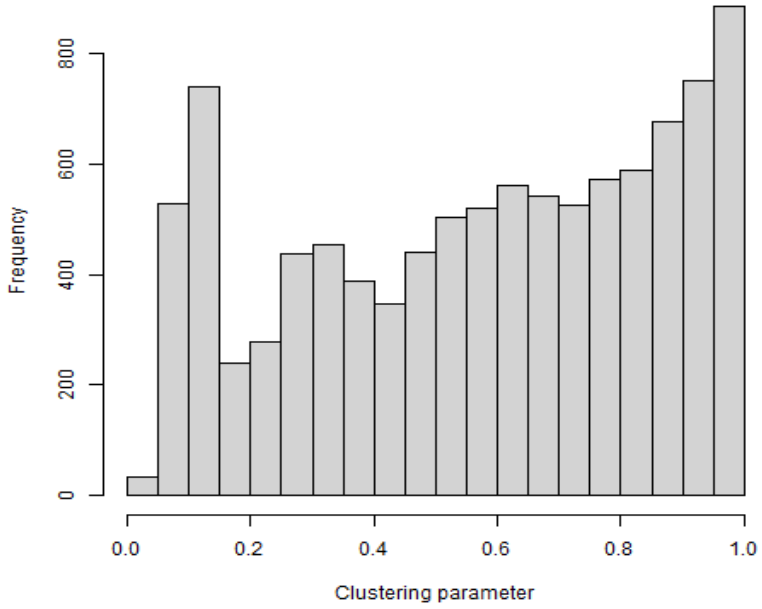


Figure 2: A histogram showing the distribution of the clustering parameter ĉ across all sampled rooms from Almeida's study.

3.2.2 The relationship between piglet-level prevalence and litter-level prevalence

Table 2: Relationship between the proportion of positive piglets in a 56-crate farrowing room and the True and Apparent (by FOF) proportion of positive litters assuming a clustering level of 0.61.

Proportion of PRRSV-positive piglets (%)	True proportion of PRRSV-positive litters (Upper and lower 95% quantiles) (%)	Apparent proportion of PRRSV-positive litters by FOF (Upper and lower 95% quantiles) (%)
1	5.36(1.79,7.14)	2.06(1.07,3.53)
5	8.93(7.14,12.50)	6.48(5.30,8.58)
10	14.29(10.71,17.86)	11.25(9.31,13.92)
15	19.64(16.07,23.21)	16.35(14.47,19.21)
20	23.21(21.43,26.79)	21.60(18.73,24.19)
25	28.57(25.00,32.14)	26.66(23.50,29.31)

I think you should include at least one paragraph to discuss the topic and to explain the tables.

Suggestion: think about what you want the reader to notice in this section. As a reader I would ask: "ok, but why did you include theses tables here? How relevant are they to your paper?

30	33.93(30.36,37.50)	31.35(28.77,34.33)
35	39.29(35.71,42.86)	36.16(33.49,39.44)
40	44.64(41.07,48.21)	41.30(38.05,44.71)
45	48.21(44.64,53.57)	46.54(43.10,49.68)
50	53.57(50.00,57.14)	51.56(48.34,54.58)

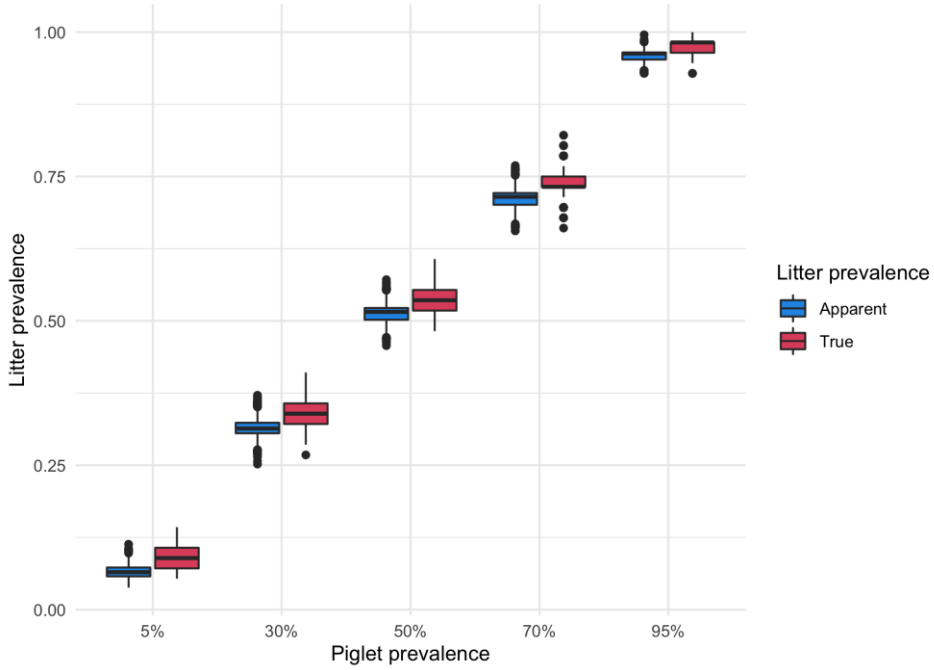


Figure 3: Distribution of True- and Apparent litter prevalence in a 56-crate room given different piglet-level prevalence scenarios and a clustering factor of 0.63.

3.3 Sensitivity analysis

The sensitivity analysis was done to evaluate the effect of variations in clustering level and room size on the proposed relationship between piglet level prevalence and litter prevalence. As can be

seen from the plots, ALP was relatively more stable to changes in clustering and the number of crates compared to TLP. Generally, TLP and ALP increasingly converged to similar values with increasing clustering and increasing room size. Clustering changes appeared to have a more significant effect on ALP and TLP than changes in the number of crates in the room (Figure 4).

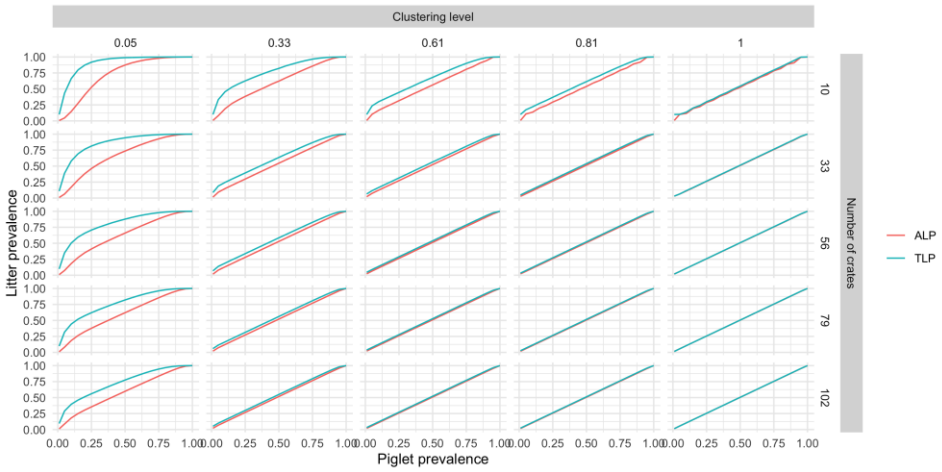


Figure 4: A graphical representation of changes in the relationship between the proportion of PRRSV-positive pigs and the proportion of PRRSV-positive litters (True and Apparent) with changes in clustering of PRRSV within room, and number of litters within rooms.

4. Discussion

The use of mathematical models to describe disease dynamics in swine populations is not new. A few examples include the use of mathematical models to characterize and describe PRRSV transmission dynamics (Amirpour Haredasht et al., 2017; Evans et al., 2010; Nodelijk et al., 2000; Phoo-ngurn et al., 2019; Rotolo et al., 2017; Suksamran et al., 2017) and in the evaluation of PRRSV control strategies (Arruda et al., 2016; Jeong et al., 2014).

Earlier studies have described the non-homogenous distribution of PRRSV in pig barns (M. N. Almeida et al., 2021; Rotolo et al., 2017). The non-homogenous aerial distribution of an infectious pathogen however is not limited to PRRSV alone (Carpenter, 2001; Kostoulas et al., 2013), and may be explained by PRRSV being more infectious than it is contagious (Pileri & Mateu, 2016), or by the mere fact that pigs in conventional US barns do not interact randomly with each other and are more likely to have the most/only contact with pigs within the same crate or with their closest neighbors (Murato et al., 2020).

Some popular statistical methods used in veterinary epidemiology for detecting and evaluating spatial (areal) clustering include Moran's *I*, ohno, black-white, Geary's *c*, and *I* pop (Carpenter, 2001), however, the use of the recursive binomial model in this study offered the authors a simplistic method to not only measure clustering, but to also propagate clustering in simulated data. The use of binomial models to detect and simulate clustering is also not new (Li et al., 2018; Nauta, 2005).

The restricted movement of pigs in conventional US swine barns and the non-homogenous distribution of viremic animals have been historically recognized to make conventional sample size assumptions (to detect a disease pathogen) not an exact fit; some previously proposed solutions include replacing simple random sampling with fixed spatial sampling (Rotolo et al., 2017), risk-based sampling (Almeida, Zhang, Zimmerman, et al., 2021), or stratified sampling (Almeida, Zhang, Lopez, et al., 2021). This study may well be one more step in trying to adjust conventional sampling schemes to better fit peculiarities with typical swine barns in the United States and with the ecology of PRRSV.

Clustering estimates the degree of homogeneity (or more aptly put; heterogeneity) of PRRSV between litters in a farrowing room; it may be overreaching to deterministically model a one-size-fits-all clustering for PRRSV. This is because the spread of PRRSV between litters within a farrowing room would depend on 1) Management practices such as cross-fostering, and vaccination (Mccaw, 2000; Pileri & Mateu, 2016) 2) PRRSV strain (there is evidence of differences in characteristics such as virulence and spread between PRRSV strains) (Cho et al., 2007; Ogno et al., 2019; Pileri & Mateu, 2016) 3) Barn structure (Rotolo et al., 2017) 4)Time since outbreak (Rotolo et al., 2017) 5) Secondary infections which may increase pig susceptibility to PRRSV, encourage huddling or increase production of infectious respiratory fluids.

The uncertainty in definitively ascertaining clustering level however does not undermine the importance of these results or pose a challenge to its utilization, on the contrary, considering/estimating clustering adds some precision to the estimated prevalence guiding sample size calculations for disease pathogen surveillance (An example is given in *Table S3*).

A critical goal of this study is to estimate the most likely relationship between the pig level prevalence and apparent litter prevalence by FOF, considering the pen-level sensitivity and specificity of this sample type. As observed from Figure 3, ALP is not as sensitive as TLP to variations in clustering parameter; we are therefore confident of the estimates on Table 2. One can also decide the number of crates or litters to randomly sample for FOF to detect PRRSV given an assumed piglet level prevalence. For example, assuming~~–~~ at least 10% piglet-level prevalence, serum sampling requires that about 30 pigs are sampled to be 95% confident of detecting at least one positive animal (Cannon & Roe, 1982; Holtkamp et al., 2011). From the table, 10% pig-level prevalence corresponds to about 11% ALP or about ~~7~~ seven litters in a 56-crate room likely to give a positive FOF test. This number can be used to calculate an appropriate sample size for FOF to detect at least ~~1~~ one positive litter; Table 2 is, therefore, useful to the practitioner in tying back what an assumed litter prevalence would be from an assumed piglet-level prevalence.

Our approach to calculating ALP implicitly considers the diagnostic performance of FOF sampling; simply put, for a given piglet-level prevalence, the difference between the ALP and TLP is due to the diagnostic performance of FOF (the probability of RT-rtPCR testing of FOF samples to correctly assign PRRSV statuses to each tested litter). This implies that ALP can be used directly to estimate FOF sample size and the only diagnostic performance we may need to consider is that of the RT-rtPCR test kit.

Another key application of the proposed tables is to help the swine practitioner estimate his piglet-level prevalence given the results of FOF testing. Given that a representative number of litters were sampled (sample size to estimate prevalence), the proportion of positive FOF results on RT-qPCR tests (apparent litter prevalence by FOF) can be used to deduce the likely proportion of viremic pigs (piglet-level prevalence).

The referenced study (Almeida, Zhang, Lopez, et al., 2021) was not designed to answer the question of spatial distribution of viremic piglets within farrowing rooms, as such, in some sampled rooms,

not every litter was sampled; consequently, the observed clustering values for those rooms may be inaccurate. To be able to deduce the number of viremic pigs from FOF positivity rate using the provided tables, it is important that one should have sampled the minimum number of litters needed to estimate prevalence.

5. Conclusion

This study explored the use of mathematical models to characterize the relationship between PP.~~,~~ TLP, and ALP in a farrowing room, this is a handy surveillance tool for weaning age pigs in typical US swine breeding herds. Another key takeaway from this study is the demonstration of the effect of a clustering parameter on the characterized relationship between the mentioned proportions; like other sampling assumptions, clustering could be considered when estimating sample size. Further similar studies on other aggregate sample types, for other subpopulations and perhaps, for other pathogens will be helpful in guiding practitioners on how they can be up-to-speed with best practice surveillance as sampling methods evolve.

REFERENCES

Almeida, M. N., Rotto, H., Schneider, P., Robb, C., Zimmerman, J. J., Holtkamp, D. J., Rademacher, C. J., & Linhares, D. C. L. (2020). Collecting oral fluid samples from due-to-wean litters. *Preventive Veterinary Medicine*, *174*. https://doi.org/10.1016/j.prevetmed.2019.104810

Almeida, M. N., Zhang, M., Lopez, W. A. L., Vilalta, C., Sanhueza, J., Corzo, C. A., Zimmerman, J. J., & Linhares, D. C. L. (2021). A comparison of three sampling approaches for detecting PRRSV in suckling piglets. *Preventive Veterinary Medicine*, *194*, 105427. https://doi.org/10.1016/J.PREVETMED.2021.105427

Almeida, M. N., Zhang, M., Zimmerman, J. J., Holtkamp, D. J., & Linhares, D. C. L. (2021). Finding PRRSV in sow herds: Family oral fluids vs. serum samples from due-to-wean pigs. *Preventive Veterinary Medicine*, *193*, 105397. https://doi.org/10.1016/j.prevetmed.2021.105397

Amirpour Haredasht, S., Polson, D., Main, R., Lee, K., Holtkamp, D., & Martínez-López, B. (2017). Modeling the spatio-temporal dynamics of porcine reproductive and respiratory syndrome cases at farm level using geographical distance and pig trade network matrices. *BMC Veterinary Research*, *13*(1). https://doi.org/10.1186/s12917-017-1076-6

Arruda, A. G., Friendship, R., Carpenter, J., Greer, A., & Poljak, Z. (2016). Evaluation of control strategies for porcine reproductive and respiratory syndrome (PRRS) in swine breeding herds using a discrete event agent-based model. *PLoS ONE*, *11*(11), 166596. https://doi.org/10.1371/journal.pone.0166596

Calderón Díaz, J. A., Fitzgerald, R. M., Shalloo, L., Rodrigues da Costa, M., Niemi, J., Leonard, F. C., Kyriazakis, I., & García Manzanilla, E. (2020). Financial Analysis of Herd Status and Vaccination Practices for Porcine Reproductive and Respiratory Syndrome Virus, Swine Influenza Virus, and Mycoplasma hyopneumoniae in Farrow-to-Finish Pig Farms Using a Bio-Economic Simulation Model. *Frontiers in Veterinary Science*, *7*, 922. https://doi.org/10.3389/FVETS.2020.556674/BIBTEX

Cameron, A. R., Meyer, A., Faverjon, C., & Mackenzie, C. (2020). Quantification of the sensitivity of early detection surveillance. *Transboundary and Emerging Diseases*, *67*(6), 2532–2543. https://doi.org/10.1111/TBED.13598

Cannon, R. M., & Roe, R. T. (1982). *Livestock disease surveys . A field manual for veterinarians. Bureau of Rural Science*, *Department of Primary Industry*. Australian Government Pub. Service. https://books.google.fr/books?id=2P6sOSdHmx0C

Carpenter, T. E. (2001). Methods to investigate spatial and temporal clustering in veterinary epidemiology. *Preventive Veterinary Medicine*, *48*(4), 303–320. https://doi.org/10.1016/S0167-5877(00)00199-9

Cho, J. G., Deen, J., & Dee, S. A. (2007). Influence of isolate pathogenicity on the aerosol transmission of Porcine reproductive and respiratory syndrome virus. *Canadian Journal of Veterinary Research*, *71*(1), 23–27.

Christopher-Hennings, J., Holler, L. D., Benfield, D. A., & Nelson, E. A. (2001). Detection and

duration of porcine reproductive and respiratory syndrome virus in semen, serum, peripheral blood mononuclear cells, and tissues from Yorkshire, Hampshire, and Landrace boars. *Journal of Veterinary Diagnostic Investigation*, *13*(2), 133–142. https://doi.org/10.1177/104063870101300207

Evans, C. M., Medley, G. F., Creasey, S. J., & Green, L. E. (2010). A stochastic mathematical model of the within-herd transmission dynamics of porcine reproductive and respiratory syndrome virus (PRRSV): Fade-out and persistence. *Preventive Veterinary Medicine*, *93*(4), 248–257. https://doi.org/10.1016/J.PREVETMED.2009.11.001

Fosgate, G. T. (2009). Practical sample size calculations for surveillance and diagnostic investigations. In *Journal of Veterinary Diagnostic Investigation* (Vol. 21, Issue 1, pp. 3–14). Journal of Veterinary Diagnostic Investigation. https://doi.org/10.1177/104063870902100102

Holtkamp, D. J., Kliebenstein, J. B., Neumann, E., Zimmerman, J. J., Rotto, H., Yoder, T. K., Wang, C., Yeske, P., Mowrer, C. L., Haley, C. A., Neumann, E. J., Rotto, H. F., & Yeske, P. E. (2013). Assessment of the economic impact of porcine reproductive and respiratory syndrome virus on United States pork producers . In *Journal of Swine Health and Production* (Vol. 21). http://lib.dr.iastate.edu/econ_las_pubs/50

Holtkamp, D. J., Polson, D. D., Torremorell, M., Morrison, B., Classen, D. M., Becton, L., Henry, S., Rodibaugh, M. T., Rowland, R. R., Snelson, H., Straw, B., Yeske, P., & Zimmerman, J. (2011). Terminology for classifying swine herds by porcine reproductive and respiratory syndrome virus status. In *Journal of Swine Health and Production* (Vol. 19, Issue 1, pp. 44–56). http://www.aasv.org/shap.html.

Holtkamp, D. J., Torremorell, M., Corzo, C. A., L Linhares, D. C., Almeida, M. N., Yeske, P., Polson, D. D., Becton, L., Snel-son, H., Donovan, T., Pittman, J., Johnson, C., Vilalta, C., Silva, G. S., & Sanhueza, J. (2021). Proposed modifications to porcine reproductive and respiratory syndrome virus herd classification. *J Swine Health Prod*, *29*(5), 261–270. http://www.aasv.org/shap.html.

Jeong, J., Aly, S. S., Cano, J. P., Polson, D., Kass, P. H., & Perez, A. M. (2014). Stochastic model

of porcine reproductive and respiratory syndrome virus control strategies on a swine farm in the United States. *American Journal of Veterinary Research*, *75*(3), 260–267. https://doi.org/10.2460/AJVR.75.3.260

Kostoulas, P., Nielsen, S. S., Browne, W. J., & Leontides, L. (2013). Sample size estimation to substantiate freedom from disease for clustered binary data with a specific risk profile. *Epidemiology and Infection*, *141*(6), 1318–1327. https://doi.org/10.1017/S0950268812001938

Li, Q., Noel-MacDonnell, J. R., Koestler, D. C., Goode, E. L., & Fridley, B. L. (2018). Subject level clustering using a negative binomial model for small transcriptomic studies. *BMC Bioinformatics*, *19*(1). https://doi.org/10.1186/s12859-018-2556-9

Linhares, D. C. L., Cano, J. P., Torremorell, M., & Morrison, R. B. (2014). Comparison of time to PRRSv-stability and production losses between two exposure programs to control PRRSv in sow herds. *Preventive Veterinary Medicine*, *116*(1–2), 111–119. https://doi.org/10.1016/J.PREVETMED.2014.05.010

Mccaw, M. B. (2000). Case report Effect of reducing crossfostering at birth on piglet mortality and performance during an acute outbreak of porcine reproductive and respiratory syndrome. *Number 1 Swine Health Prod*, *8*(1), 15–21. http://www.aasp.org/shap.html.

Murato, Y., Hayama, Y., Shimizu, Y., Sawai, K., & Yamamoto, T. (2020). *Evaluation of sampling methods for effective detection of infected pig farms during a disease outbreak.* https://doi.org/10.1371/journal.pone.0241177

Nauta, M. J. (2005). Microbiological risk assessment models for partitioning and mixing during food handling. *International Journal of Food Microbiology*, *100*(1–3), 311–322. https://doi.org/10.1016/j.ijfoodmicro.2004.10.027

Nodelijk, G., De Jong, M. C. M., Van Nes, A., Vernooy, J. C. M., Van Leengoed, L. A. M. G., Pol, J. M. A., & Verheijden, J. H. M. (2000). Introduction, persistence and fade-out of porcine reproductive and respiratory syndrome virus in a Dutch breeding herd: A mathematical analysis. *Epidemiology and Infection*, *124*(1), 173–182.

https://doi.org/10.1017/S0950268899003246

Nunes de Almeida, M. (2020). Improved Porcine Reproductive and Respiratory Syndrome Virus Surveillance in Swine Breeding Herds [Iowa State University]. In *ProQuest Dissertations and Theses*. https://doi.org/10.31274/etd-20210114-104

Ogno, G., Rodríguez-Gómez, I. M., Canelli, E., Ruedas-Torres, I., Álvarez, B., Domínguez, J., Borghetti, P., Martelli, P., & Gómez-Laguna, J. (2019). Impact of PRRSV strains of different in vivo virulence on the macrophage population of the thymus. *Veterinary Microbiology*, *232*, 137–145. https://doi.org/10.1016/j.vetmic.2019.04.016

Osemeke, O. H., de Freitas Costa, E., Almeida, M. N., Trevisan, G., Ghosh, A., Silva, G. S., & Linhares, D. C. L. (2022). Effect of pooling family oral fluids on the probability of PRRSV RNA detection by RT-rtPCR. *Preventive Veterinary Medicine*, 105701. https://doi.org/10.1016/J.PREVETMED.2022.105701

Phoo-ngurn, P., Kiataramkul, C., & Chamchod, F. (2019). Modeling the spread of porcine reproductive and respiratory syndrome virus (PRRSV) in a swine population: transmission dynamics, immunity information, and optimal control strategies. *Advances in Difference Equations*, *2019*(1), 1–12. https://doi.org/10.1186/S13662-019-2351-6/FIGURES/5

Pileri, E., & Mateu, E. (2016). Review on the transmission porcine reproductive and respiratory syndrome virus between pigs and farms and impact on vaccination. In *Veterinary Research* (Vol. 47, Issue 1, pp. 1–13). BioMed Central. https://doi.org/10.1186/s13567-016-0391-4

Rotolo, M. L., Sun, Y., Wang, C., Giménez-Lirola, L., Baum, D. H., Gauger, P. C., Harmon, K. M., Hoogland, M., Main, R., & Zimmerman, J. J. (2017). Sampling guidelines for oral fluid-based surveys of group-housed animals. *Veterinary Microbiology*, *209*, 20–29. https://doi.org/10.1016/j.vetmic.2017.02.004

Rovira, A., Reicks, D., & Muñoz-Zanzi, C. (2007). Evaluation of surveillance protocols for detecting porcine reproductive and respiratory syndrome virus infection in boar studs by simulation modeling. *Journal of Veterinary Diagnostic Investigation*, *19*(5), 492–501. https://doi.org/10.1177/104063870701900506

Silva, G. S., Schwartz, M., Morrison, R. B., & Linhares, D. C. L. (2017). Monitoring breeding herd production data to detect PRRSV outbreaks. *Preventive Veterinary Medicine*, *148*, 89–93. https://doi.org/10.1016/J.PREVETMED.2017.10.012

Stevenson, M. A. (2021). Sample Size Estimation in Veterinary Epidemiologic Research. In *Frontiers in Veterinary Science* (Vol. 7, p. 539573). Frontiers Media S.A. https://doi.org/10.3389/fvets.2020.539573

Suksamran, J., Lenbury, Y., Satiracoo, P., & Rattanakul, C. (2017). A model for porcine reproductive and respiratory syndrome with time-dependent infection rate: traveling wave solution. *Adv Differ Equ*, *2017*, 215. https://doi.org/10.1186/s13662-017-1282-3

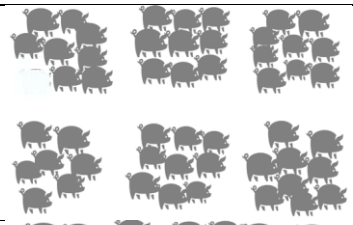
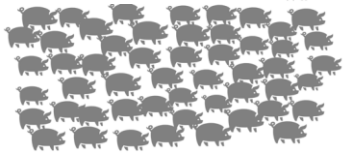
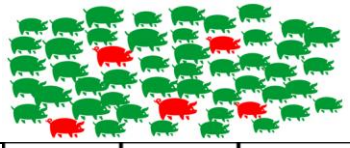
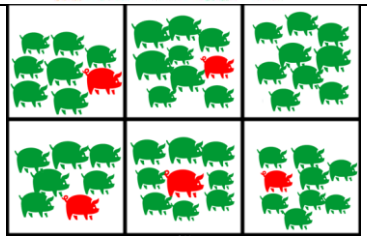
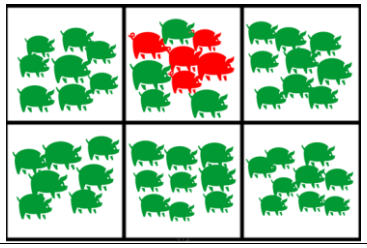
Trevisan, G., Linhares, L. C. M., Crim, B., Dubey, P., Schwartz, K. J., Burrough, E. R., Main, R. G., Sundberg, P., Thurn, M., Lages, P. T. F., Corzo, C. A., Torrison, J., Henningson, J., Herrman, E., Hanzlicek, G. A., Raghavan, R., Marthaler, D., Greseth, J., Clement, T., … Linhares, D. C. L. (2019). Macroepidemiological aspects of porcine reproductive and respiratory syndrome virus detection by major United States veterinary diagnostic laboratories over time, age group, and specimen. *PLoS ONE*, *14*(10), e0223544. https://doi.org/10.1371/journal.pone.0223544

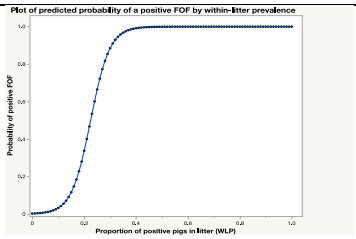
Turlewicz-Podbielska, H., Włodarek, J., & Pomorska-Mól, M. (2020). Noninvasive strategies for surveillance of swine viral diseases: a review. *Journal of Veterinary Diagnostic Investigation : Official Publication of the American Association of Veterinary Laboratory Diagnosticians, Inc*, *32*(4), 503. https://doi.org/10.1177/1040638720936616

SUPPLEMENTARY MATERIALS

Figure S5: A general description of the stochastic model with pictorial illustrations

	Goal	Example	Example in pictures
1	**To create a farrowing room with *n* litters** by generating *n*	From this empirical distribution, a	

	random numbers from a discrete empirical distribution corresponding to the average litter size we observed from actual field data	possible set of random numbers generated will be 8, 9, 9, 7, 9, 8. These numbers are shown in the picture on the right.	
2.	**To create a disease prevalence scenario within this created room**.	When we add all the litters (generated numbers), we have 52 pigs (8+9+9+7+9+8 = 50).	
	Suppose we wanted a 10% prevalence. We will want 10% of the total pigs in the simulated room to be positive.	10% of 50 is 5, meaning five pigs will be PRRSV positive	
3	**To distribute the diseased animals between pens in a manner typical of PRRSV**. PRRSV has been repeatedly reported to be heterogeneously distributed (clustered) within a farrowing room. A clustering factor in the recursive binomial model, which could range between 0 (homogenous distribution) and 1 (complete clustering), was used to assign positive pigs to litters. The clustering factor determines the True Litter Prevalence (TLP), defined as the number of litters with at least one positive pig.	The first image shows what would be obtainable if there was no clustering (Clustering = 0). **TLP** = 5/6	
		The next image shows what is expected in complete clustering (clustering = 1). **TLP** = 1/6.	

4	**To determine the expected number of positive FOF from this room if all the litters are tested**. A predictive model is fitted using data from a previous study (Almeida, Zhang, Lopez, et al., 2021). This model gives the probability of a positive FOF (pFOF) test given the proportion of positive pigs within a litter or within-litter prevalence (WLP).	The graph shows the relationship between pFOF and WLP.	
		For the clustering = 0, the WLP and pFOF for each litter are calculated. The expected number of positive FOF for that room is 0.2998 (<1). **ALP** = <1/6	WLP = 0.125 pFOF = 0.062 | WLP = 0.111 pFOF = 0.043 | WLP = 0 pFOF = 0.002 WLP = 0.143 pFOF = 0.099 | WLP = 0.111 pFOF = 0.043 | WLP = 0.111 pFOF = 0.043
	After these probabilities (pFOFs) are generated, the Expected number of positive FOF tests for that room is determined using a Monte Carlo process. The Apparent litter prevalence by FOF (ALP) is then the expected number of positive FOF divided by the total number of litters or tests	For the clustering = 1. The expected number of positive FOF for that room is 1.014 (approximately 1). **ALP** = 1/6	WLP = 0 pFOF = 0.002 | WLP = 0.556 pFOF = 1.000 | WLP = 0 pFOF = 0.002 WLP = 0 pFOF = 0.002 | WLP = 0 pFOF = 0.002 | WLP = 0 pFOF = 0.002
5 6	**Repeat steps 1 to 4 (4,999 more times)** and obtain the median TLP and ALP thereafter. The 5% prevalence is then matched with the median TLP and median ALP for the chosen clustering level. Then repeat steps 1 to 5 for other prevalence scenarios		

437

438

439 *Table 3: Relationship between PP, ALP, and TLP at different clustering levels and room sizes*

PP (%)	Number of crates (%)	Clustering (%)	ALP (%)	TLP (%)
1.00	10	5.00	0.43	10.00
5.00	10	5.00	3.55	40.00
10.00	10	5.00	14.24	70.00
15.00	10	5.00	27.20	80.00
20.00	10	5.00	40.68	90.00
25.00	10	5.00	52.83	90.00
30.00	10	5.00	63.14	100.00
35.00	10	5.00	71.21	100.00
40.00	10	5.00	78.01	100.00
45.00	10	5.00	83.62	100.00
50.00	10	5.00	87.88	100.00
1.00	10	33.00	0.45	10.00
5.00	10	33.00	8.48	30.00
10.00	10	33.00	19.49	50.00
15.00	10	33.00	27.50	50.00
20.00	10	33.00	32.27	60.00
25.00	10	33.00	38.56	60.00
30.00	10	33.00	42.51	70.00
35.00	10	33.00	48.40	70.00
40.00	10	33.00	52.21	70.00
45.00	10	33.00	58.21	80.00
50.00	10	33.00	61.85	80.00
1.00	10	61.00	0.45	10.00
5.00	10	61.00	10.31	20.00
10.00	10	61.00	17.35	30.00
15.00	10	61.00	20.87	30.00
20.00	10	61.00	27.25	40.00
25.00	10	61.00	30.81	40.00
30.00	10	61.00	37.26	50.00
35.00	10	61.00	40.65	50.00
40.00	10	61.00	46.87	60.00
45.00	10	61.00	50.57	60.00
50.00	10	61.00	56.75	70.00
1.00	10	100.00	0.46	10.00
5.00	10	100.00	10.19	10.00
10.00	10	100.00	10.23	10.00

15.00	10	100.00	20.15	20.00
20.00	10	100.00	20.22	20.00
25.00	10	100.00	30.13	30.00
30.00	10	100.00	30.19	30.00
35.00	10	100.00	40.11	40.00
40.00	10	100.00	40.17	40.00
45.00	10	100.00	50.09	50.00
50.00	10	100.00	50.14	50.00
1.00	56	5.00	0.64	10.71
5.00	56	5.00	7.35	35.71
10.00	56	5.00	18.28	50.00
15.00	56	5.00	27.65	58.93
20.00	56	5.00	34.97	66.07
25.00	56	5.00	41.26	71.43
30.00	56	5.00	46.64	75.00
35.00	56	5.00	51.69	78.57
40.00	56	5.00	56.51	82.14
45.00	56	5.00	61.27	85.71
50.00	56	5.00	66.01	87.50
1.00	56	33.00	1.75	7.14
5.00	56	33.00	7.92	14.29
10.00	56	33.00	13.02	19.64
15.00	56	33.00	18.04	23.21
20.00	56	33.00	23.13	28.57
25.00	56	33.00	27.93	33.93
30.00	56	33.00	32.86	39.29
35.00	56	33.00	37.92	42.86
40.00	56	33.00	42.97	48.21
45.00	56	33.00	47.89	53.57
50.00	56	33.00	52.86	58.93
1.00	56	61.00	2.06	5.36
5.00	56	61.00	6.45	8.93
10.00	56	61.00	11.27	14.29
15.00	56	61.00	16.34	19.64
20.00	56	61.00	21.61	23.21
25.00	56	61.00	26.62	28.57
30.00	56	61.00	31.36	33.93
35.00	56	61.00	36.18	39.29
40.00	56	61.00	41.28	44.64
45.00	56	61.00	46.54	48.21

50.00	56	61.00	51.62	53.57
1.00	56	100.00	2.00	1.79
5.00	56	100.00	5.57	5.36
10.00	56	100.00	10.91	10.71
15.00	56	100.00	16.23	16.07
20.00	56	100.00	20.42	21.43
25.00	56	100.00	25.18	25.00
30.00	56	100.00	30.52	30.36
35.00	56	100.00	35.85	35.71
40.00	56	100.00	41.18	41.07
45.00	56	100.00	45.19	46.43
50.00	56	100.00	50.13	50.00
1.00	102	5.00	0.91	9.80
5.00	102	5.00	8.75	29.41
10.00	102	5.00	18.14	39.22
15.00	102	5.00	24.96	46.08
20.00	102	5.00	30.44	50.98
25.00	102	5.00	35.47	55.88
30.00	102	5.00	40.36	60.78
35.00	102	5.00	45.31	64.71
40.00	102	5.00	50.15	68.63
45.00	102	5.00	54.99	73.53
50.00	102	5.00	59.81	77.45
1.00	102	33.00	2.07	4.90
5.00	102	33.00	6.85	9.80
10.00	102	33.00	11.85	14.71
15.00	102	33.00	16.82	19.61
20.00	102	33.00	21.76	24.51
25.00	102	33.00	26.75	30.39
30.00	102	33.00	31.67	35.29
35.00	102	33.00	36.66	40.20
40.00	102	33.00	41.68	45.10
45.00	102	33.00	46.69	50.00
50.00	102	33.00	51.68	54.90
1.00	102	61.00	1.99	2.94
5.00	102	61.00	6.04	6.86
10.00	102	61.00	10.99	11.76
15.00	102	61.00	15.90	17.65
20.00	102	61.00	20.82	22.55
25.00	102	61.00	25.77	27.45

30.00	102	61.00	30.74	32.35
35.00	102	61.00	35.70	37.25
40.00	102	61.00	40.81	42.16
45.00	102	61.00	45.89	47.06
50.00	102	61.00	50.88	51.96
1.00	102	100.00	1.21	0.98
5.00	102	100.00	5.14	5.88
10.00	102	100.00	10.31	10.78
15.00	102	100.00	15.68	15.69
20.00	102	100.00	20.74	20.59
25.00	102	100.00	25.65	25.49
30.00	102	100.00	30.54	30.39
35.00	102	100.00	35.44	35.29
40.00	102	100.00	40.33	40.20
45.00	102	100.00	45.23	45.10
50.00	102	100.00	50.12	50.00